BUBBLEGUM RESEARCH INSTITUTE
BRI
OUR MISSION:
TO UNDERSTAND THE WORLD, ONE BUBBLE AT A TIME
Best Average Linear Mapping
For a layer ℓ and context c:
$W_c^\ell = \arg\min_W \mathbb{E}_{(x,y)\sim D}[\|Wx - y\|^2 \mid c]$
Interpretation:
The linear mapping that minimizes average error for this context at this layer.
Context c is determined by the input representation.
Similar representations activate this context
Local. Average. Linear. Powerful.
TODAY'S QUESTION:
HOW DOES CONTEXT SHAPE COMPUTATION?
SCIENCE IS SWEETER WITH GUM
LINEAR ALGEBRA FOR GUM SCIENTISTS
PROBABILITY AND BUBBLES
STATISTICS CHEWS
MATRIX METHODS MADE SWEET
Key Idea:
Context → Selects Mapping
Mapping → Transforms Representation
Goal:
Learn the best average linear mapping for every context at every layer.
RESEARCH. CHEW. REPEAT.
STRAWBERRY
BLUEBERRY
ORANGE MARMALADE
COTTON CANDY
CLASSIC BUBBLEGUM
TODAY'S GUM TASTING!

Best average linear mapping per context per layer for a best total composition of linear mappings over the network.